The first paper “Developing a Data-Driven Method for Estimating Provider Penetration and Abusive Billing Practices”, addresses the task of identifying providers that appear to be generating more claims than expected given patient and provider attributes. A provider’s penetration into a market is a key variable in monitoring for suspicious behavior, and a two-stage model uses clustering to identify the provider’s market area and HCFA data sources to model provider penetration.

The second paper, “Patterns Extraction for Monitoring Medical Practices,” investigates application of relational patterns to the monitoring of hospital discharge abstracts for quality of care. They demonstrate application of relational patterns to identify poor hospitalization practices.

The third paper “Precursory Steps to Mining HCFA Health Care Claims” discusses the work that must be performed prior to the actual data mining. These tasks include: customer discussions, data extraction and cleaning, transformation of the database, and auditing (basic statistics and visualization of the information) of the data.

The fourth paper, “A Global Optimization Approach to Cluster Analysis in Medical Diagnosis and Prognosis.” applies a technique based on convex and global optimization to Breast Cancer databases and achieves improvements in diagnosis and prognosis prediction over previously reported results.

The fifth paper analyzes data from the Health Insurance Commission (the Australian equivalent of HCFA). In “Data Mining of Administrative Claims Data for Pathology Services,” a number of new features are identified for use in predictive modeling. These features are summarized, visualized and used as inputs for clustering and outlier detection methods.

The sixth paper, presents “Descriptive Modeling in Healthcare Supporting a Facility Location Decision via GIS-Based Market Visualization.” This paper illustrates the decision support power of combining publicly available and system-specific data with a Geographic Information System (GIS) for locating and sizing a proposed Neonatal Intensive Care Unit (NICU) within a system’s network of rural hospitals.

Finally, the seventh paper, “Empirical Norms as a Lever for On-line Support of General Practice,” explores improvement of data entry efficiency and practice quality in primary care systems. The paper applies Bayesian models to a large primary care database to produce adaptive menus and dynamic decision support alerts.

This diverse selection of papers demonstrates the range of possibilities in data mining for healthcare quality, efficiency and practice support. It shows the range of application timing from point-of-care to financial review to facilities planning. It shows the range of methods from graphical display to novel query pattern and cluster analysis techniques. In all cases, the papers leave us with the realization that the healthcare systems can be significantly improved through the continued improvement and use of these techniques. Clearly there is scope for a return visit to Hawaii in 2002!